IET Journals
The Institution of Engineering and Technology

Automation enhancement and accuracy investigation of a portable single-camera gait analysis system

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

C. Yang[1,2,3] ✉ *U. C. Ugbolue*[3,4] *D. McNicol*[3] *V. Stankovic*[3] *L. Stankovic*[3] *A. Kerr*[5] *B. Carse*[5] *K. T. Kaliarntas*[6] *P. J. Rowe*[5]

[1] *Institute of Computer Vision, College of Computer Science and Technology, Zhejiang University of Technology, HangZhou 310014, China*
[2] *Digital Content and Media Sciences Research Division, National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
[3] *Department of Electronic and Electrical Engineering, University of Strathclyde, Royal College Building, 204 George St, Glasgow G1 1XW, UK*
[4] *Institute of Clinical Exercise & Health Science, University of the West of Scotland, Lanarkshire Campus, South Lanarkshire G72 0LH, UK*
[5] *Department of Biomedical Engineering, University of Strathclyde, Wolfson Building, 106 Rottenrow, Glasgow G4 0NW, UK*
[6] *School of Applied Sciences, Department of Sport, Exercise & Health Science, Edinburgh Napier University, Edinburgh EH11 4BN, UK*
✉ *E-mail: cheng@nii.ac.jp*

Abstract: While optical motion analysis systems can provide high-fidelity gait parameters, they are usually impractical for local clinics and home use, due to high cost, requirement for large space, and lack of portability. In this study, we focus on a cost-effective and portable, single-camera gait analysis solution, based on video acquisition with calibration, autonomous detection of frames-of-interest, Kalman-filter+Structural-Similarity-based marker tracking, and autonomous knee angle calculation. The proposed system is tested using 15 participants, including 10 stroke patients and 5 healthy volunteers. The evaluation of autonomous frames-of-interest detection shows only 0.2% difference between the frame number of the detected frame compared to the frame number of the manually labelled ground truth frame, and thus can replace manual labelling. The system is validated against a gold standard optical motion analysis system, using knee angle accuracy as metric of assessment. The accuracy investigation between the RGB- and the grayscale-video marker tracking schemes shows that the grayscale system suffers from negligible accuracy loss with a significant processing speed advantage. Experimental results demonstrate that the proposed system can automatically estimate the knee angle, with R-squared value larger than 0.95 and Bland-Altman plot results smaller than 3.0127 degrees mean error.

Notation

Symbol	Property
$\mathbf{D}$	a video file
F	a video frame
N	total number of frames in a video
n	frame number
S	frame-segment
H	histogram
M	total number of quantization bins of the histogram
m	quantization bin
d	total number of the pixels at the same quantization bin
P	detected peaks
$\hat{\mathbf{s}}^{-}$	a priori estimate
$\hat{\mathbf{s}}$	a posteriori estimate
$\hat{\mathbf{u}}$	centre coordinate of the marker
$\mathbf{R}$	state transition matrix
$\mathbf{X}^{-}$	a posteriori covariance matrix
$\mathbf{B}$	process noise covariance matrix
$\mathbf{K}$	Kalman gain
$\mathbf{Q}$	observation matrix
$\mathbf{E}$	measurement error covariance matrix
f	column centre coordinate of the search area
g	row centre coordinate of the search area
v	K-velocity
c	column centre coordinate of the marker
r	row centre coordinate of the marker
t	duration of one frame
$\mathbf{a}$	candidate block
$\mathbf{b}$	marker-template
$\in$	set membership,

1 Introduction

Advanced objective clinical gait analysis on stroke patients can generate quantified, standardised, and more reliable gait measurements [1] compared to traditional, semi-subjective [2], observational gait analysis methods [3, 4], while being minimally intrusive to the stroke patients [2, 5]. Some examples include acoustic gait analysis systems [6], optical non-wearable motion analysis systems [2, 7] based on strategically located infrared cameras to capture three dimensional (3D) limb motion by tracking retroreflective markers adhered to the skin overlying anatomical landmarks of the study participants, and markerless systems that are completely contact-less to patients, such as Organic Motion OpenStage 2.0 (Organic Motion HQ, New York, NY). However, all these systems have downsides, such as: (1) they require operational expertise and large laboratory space, hence patients need to be regularly transported to major clinics for assessment; (2) they do not facilitate easy comparison with results from a previous assessment in a longitudinal study; (3) they are expensive; and (4) they cannot distinguish between gradual and abrupt functional changes which negatively affect clinical intervention [8]. Additionally, markerless systems are particularly sensitive to the motion capture background and ambient lighting, which could make patients uncomfortable. For example, OpenStage 2.0 requires white fabric walls and strong stage lights.

Motivated by cost and providing a convenient option to patients and health services, research and development on cost effective and portable systems has emerged. Related range sensor and home-video based systems, which cost about £700, such as [9], [10] and [11], that build on the work of [12], with Pro-Trainer motion analysis software (Sports Motion, Inc., Cardiff, CA), offer gait analysis outside the gait laboratory, *e.g.*, in local clinics and at homes. Similar to other range sensor and home-video based gait analysis systems [2, 13–21] and Inertial Measurement Unit (IMU) based gait analysis systems

[22–26], the gait parameters obtained after data processing can be sent to physiatrists for clinical consultation, indicating the potential for tele-rehabilitation [27–31]. It is shown in [32] that a 2D video tracker software provides similar accuracy to VICON 3D system for knee angle measurement but not for measurement of the ankle angle over time.

Motivated by the opportunity offered by [9], [11] for cost-effective tele-rehabilitation, in conference versions of this work [33, 34], we designed a portable, single-camera gait analysis system, that tracks bulls-eye markers (see Figure 1(a)) attached to the patient joints and displays the calculated joint angles. The system consists of marker tracking, based on a Kalman filter [35, 36] and Structural-Similarity [37], and autonomous knee angle calculation. In [38], we extended the system to autonomous gait event detection. Experimental results show high detection rates for all six gait events / phases in each gait cycle: Initial Contact, Foot Flat, Mid-Stance, Heel Raise, Terminal Contact, and Mid-Swing [39], in contrast to kinetics-based gait event detection methods that use force plates [39], as is conventional for most optical motion analysis systems. Though our previous system [33, 34, 38] provided a cost effective alternative to expensive and bulky optical motion analysis systems, it still required significant manual effort to operate. Moreover, it was not validated in a clinical study.

In this paper, we further enhance our initial single-camera system [33, 34, 38, 40] by enabling autonomous detection of the frame when tracking starts and ends, and investigate the performance of RGB-video vs. grayscale-video marker tracking schemes. We validate the proposed system against VICON in terms of the knee angle performance using 15 participants including stroke patients and healthy volunteers.

Our system addresses some of the drawbacks of related range sensor and home-video based systems [2, 9, 11, 13–19, 41] and IMU systems [22–26] namely: (1) Unlike [9], there are no colour restrictions on the background or the participant's clothing; (2) In contrast to Soda et al. [9], which is validated on only one healthy volunteer with one walking trial with no gold standard benchmark, we validate our proposed system's knee angle against the gold standard VICON MX Giganet 6xT40 and 6xT160 (VICON Motion Systems Ltd., Oxford, UK, approximately £250,000) optical motion analysis system (the same gold standard as used by [11]). (3) Unlike systems of [11] and Pro-Trainer and Siliconcoach (Siliconcoach Ltd., Dunedin, New Zealand) as used by [42] and [43] that require significant manual effort, our system autonomously tracks the markers attached to the joints and calculates the knee angle; the only operational effort required is for marker-template selection for tracking initialization which is done via a user-friendly graphical user interface (GUI). (4) Unlike the passive marker system [41] that is only validated on one side of the body without any benchmarking systems, our system is validated on both sides of the body with a gold standard VICON optical motion analysis system. (5) 3D Kinect range sensor-based systems [13–18, 20, 21] cannot reliably capture relatively fast body motion, since Kinect operates at only 30 frames per second (fps), whereas our system operates at 210 fps. (6) Like other range sensor and home-video based systems, our system is non-intrusive to the participants, which is in contrast to state-of-the-art IMU gait analysis systems [22–26]. However, with only a 2D camera in our gait analysis system, its drawback lies in the following two aspects: (1) Estimation of the human joint locations using our system is less accurate compared to 3D Kinect-based range sensor systems, and (2) The gait parameters derived from the 2D images in our system are less reliable than those derived from the inertial data in IMU systems.

Overall, the system is simple to assemble, highly adjustable for camera view, cost effective, and transportable for efficient gait analysis at local clinics and homes. In addition, the gait analysis result from the proposed system can be immediately sent to physicians for clinical consultation, indicating the potential to facilitate tele-rehabilitation [27–31]. These are in contrast to laboratory-based optical motion analysis systems that require large laboratory space, operational expertise, and have lots of pieces of equipment to assemble. Furthermore, our proposed system is more practical than recent single-camera approaches, that require either substantial manual effort for joint angles [11, 42, 43], or specific limitations on the video capturing background, ambient light, and clothing [9].

Note that, the proposed system is also capable of automatically measuring global segment orientations in the sagittal plane, *e.g.*, shank-to-vertical and thigh-to-vertical angles, with added bulls-eye markers at the femur and tibia, showing potential to facilitate ankle-foot orthosis fitting and tuning [44–46].

Our proposed system is more suitable for gait analysis in a rehabilitation context, providing users with feedback on kinematic changes. However, highly-accurate 3D optical motion analysis systems may still be needed to inform clinical decision making, *e.g.*, before orthopaedic surgery.

2 Method

This section provides an overview of the proposed system and describes its main building blocks.

2.1 System overview

The proposed system comprises a digital camera EX-FH20 EXILIM (Casio Computer Co., Ltd., Tokyo, Japan) with a tripod, and 6 bulls-eye black-and-white paper markers, as used in [11]. The system also uses a 10×7 calibration checkerboard [47] with square size of 23.3mm, as shown in Figure 1(b), and a laptop with bespoke data processing software developed in Matlab R2014b (MathWorks, Inc., Natick, MA). The system goal is to autonomously analyze the study participant's gait patterns indicated by knee angle.

Before video acquisition, 6 bulls-eye markers (see Figure 1(a)), each with an outside diameter of 4.5cm and an inner diameter of 2.2cm, are attached to the skin overlaying the hip, knee, and ankle joint centres on the external part of both legs of a study participant in the sagittal plane. As shown in Figure 1(c), the study participants walk from left to right, and back on a 6m×0.8m mat using a similar approach to [11]. The digital camera is configured at 360×480 pixel resolution, 210 frames per second (fps), mounted on a tripod to a height of 0.5-1.0m and positioned 2.0m away from the long-side centre of the mat, depending on the study participant, and calibrated using [47] with the checkerboard (see Figure 1(b)), including calibration board feature point detection, camera intrinsic and extrinsic parameter estimation using a closed-form solution and radial distortion coefficient estimation using nonlinear minimization, to remove lens distortion in the video frames, which are then processed for marker tracking and knee angle calculation.

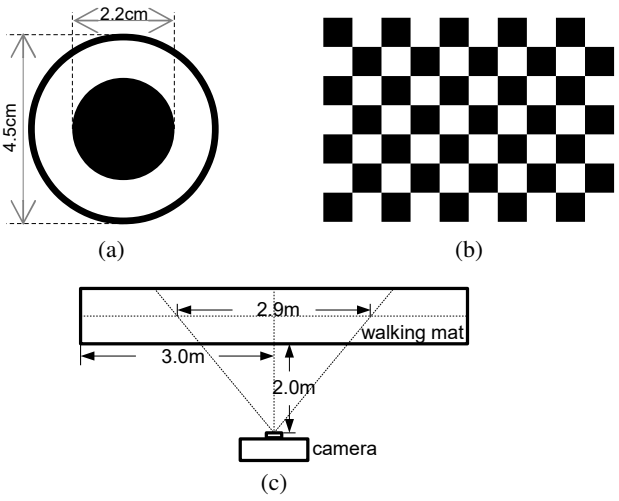


Fig. 1: (a) A sample black-and-white bulls-eye marker; (b) Calibration checkerboard; (c) Proposed single-camera gait analysis system.

2.2 *Video acquisition*

For benchmarking with the gold standard VICON system, the SWIFT Cast trial protocol [48, 49] is applied in the stroke-patient group and Plug-in-Gait protocol [50] in the healthy-volunteer group. For each stroke patient, retroreflective markers (14mm-diameter) are fixed to the skin overlaying the anatomical landmarks, as done in [49]. The knee flexion / extension axes are determined based on marker clusters at the femur and tibia and single calibration markers, followed by corresponding joint angle calculation as in [49]. For each healthy volunteer, 15 retroreflective markers (14mm in diameter) are fixed to the skin overlying the following anatomical landmarks adapted from the Plug-in-Gait protocol [50], denoted as: sacral wand marker, left (right) anterior superior iliac spine, knee, femur, ankle, tibia, toe, and heel markers, followed by joint angle calculation based on the Euler / Cardan angle determination algorithm with an *y-x-z* axis rotation sequence, namely flexion / extension, adduction / abduction, and internal / external rotation [50]. For both groups, all VICON motion-capture modules are calibrated. Note that, both bulls-eye and retroreflective markers are attached onto the knee and ankle joints, where each bulls-eye marker is firstly attached onto the corresponding anatomical landmark and the retroreflective marker is attached on top of that bulls-eye marker. Such marker placement does not adversely affect marker tracking performance. Specifically, for VICON, the thickness of the paper-made, bulls-eye marker is negligible, and therefore there is negligible error in joint centre calculation based on the retroreflective markers; for our system, adhering retroreflective markers on top of bulls-eye markers only changes the appearance of the region of interest for marker tracking, which does not introduce additional source of error. We will discuss the effect of the above marker placement in detail in Section 4.

Each study participant is simultaneously recorded using the proposed system and VICON. Figure 2(a) shows a sample single-camera scene for a healthy volunteer, where 4 out of 12 VICON infrared cameras are marked with red circles and 3 bulls-eye markers on the left leg of the study participant are marked with yellow circles.

2.3 *Autonomous frames-of-interest detection*

Video recording starts when the study participant begins walking even though he/she is still not within the camera's field of view. The method needs to automatically recognize the first and last frames when all three markers are present (called "frames-of-interest" (FOI)) to start the marker tracking process. Due to a noticeable change of the frame histogram when a study participant walks into and out of the camera scene, we propose an image histogram-based frames-of-interest detection scheme that identifies at which frame the system starts tracking the markers and at which frame tracking stops. An image histogram shows the number of pixels of each intensity in a frame. To recognize entrance from both left and right side of the field of view, we define two frame segments as shown in Figure 2(a), denoted as S_1 and S_2, for all N frames, denoted as $\{\mathrm{F}_1, ..., \mathrm{F}_N\}$, in a video file.

For each frame segment of the N frames, we compute the histograms, denoted as $\{\mathrm{H}_1^x, ..., \mathrm{H}_N^x\}$, where $x \in \{\mathrm{S}_1, \mathrm{S}_2\}$ and for each $\mathrm{H}_n^x \in \{\mathrm{H}_1^x, ..., \mathrm{H}_N^x\}$, $\mathrm{H}_n^x = \{d_{n1}^x, ..., d_{nm}^x\}$, where n is the frame number, $1 \le n \le N$, $1 \le m \le M$. M denotes the number of quantization bins of each histogram, and d_m is the total number of pixels in the m-th quantization bin. Next, we compute the difference between H_n^x and H_1^x, for all n, denoted as $\Delta\mathrm{H}_n^x$: $\Delta\mathrm{H}_n^x = \{|d_{n1}^x - d_{11}^x|, ..., |d_{nm}^x - d_{1m}^x|\}$, followed by forming the element sum of $\Delta\mathrm{H}_n^x$, $\sum_{j=1}^{m} |d_{nj}^x - d_{1j}^x|$, denoted as $\sum \Delta\mathrm{H}_n^x$. The $\{n, \sum \Delta\mathrm{H}_n^x\}$ plot of a trial that contains walking from right to left and then from left to right is shown in Figure 2(b). We then perform peak detection in the $\{n, \sum \Delta\mathrm{H}_n^x\}$ plot for each frame segment with a heuristically set threshold τ, where the detected peaks with corresponding frame numbers are denoted as P^x: $\mathrm{P}^x = \{l^x, \sum \mathrm{H}_{l^x}^x\}$, and $\forall l^x$ ($\forall$ denotes universal quantification), $\{\sum \Delta\mathrm{H}_{l^x}^x \ge \tau\}$. The detected peaks are marked with red asterisks in Figure 2(b).

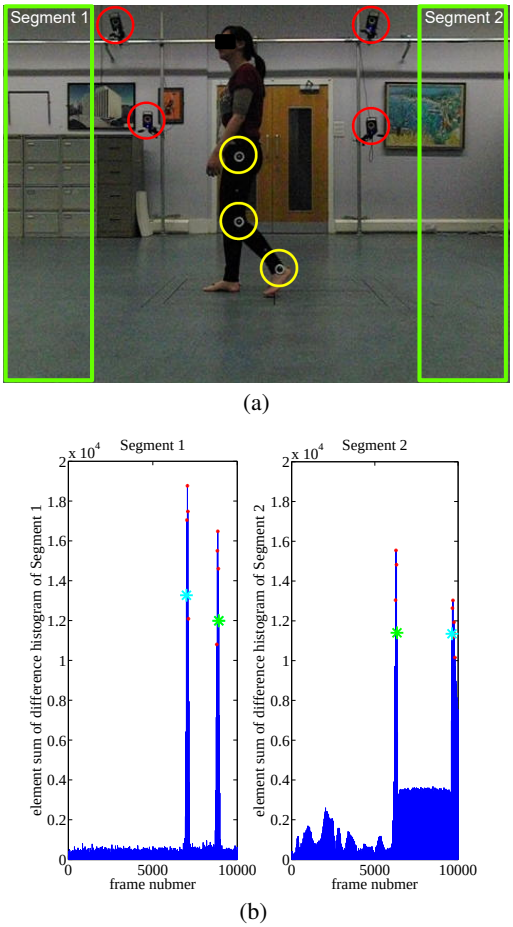


Fig. 2: (a) Sample video frame when the study participant is walking into the camera scene with two frame-segments for the detection of the "frames-of-interest"; (b) The $\{n, \sum \Delta\mathrm{H}_n^x\}$ plots for two frame-segments with marked detected first (green asterisks) and last (cyan asterisks) frames of interest.

Each video file $\mathbf{D}$ contains a pair of left-to-right (LtR) and right-to-left (RtL) walking trials – a study participant walks into and out of the camera scene twice, once from each direction, indicating that there exist two peak clusters, as shown in Figure 2(b). To separate P^x into two clusters, we first compute the difference of frame numbers between neighbouring peaks in P^x, denoted as Δl^x, with $\Delta l_{\max}^x = \arg\max_{l^x} \Delta l^x$, and let l_{m1}^x, l_{m2}^x be the two corresponding frame numbers, i.e., $\Delta l_{\max}^x = l_{m2}^x - l_{m1}^x$. Then we separate P^x as follows:

$\mathrm{P}_1^x = \{\text{a subset of } \{l^x, \sum \mathrm{H}_{l^x}^x\},\ l^x \le l_{m1}^x\}$;
$\mathrm{P}_2^x = \{\text{a subset of } \{l^x, \sum \mathrm{H}_{l^x}^x\},\ l_{m2}^x \le l^x\}$.

Let the frame numbers associated with the first and last detected peaks in x be l_{first}^x and l_{last}^x, respectively. If $\Delta l_{\max}^{\mathrm{S}_1} > \Delta l_{\max}^{\mathrm{S}_2}$, which indicates the trial directions in $\mathbf{D}$ are LtR $\rightarrow$ RtL, we designate the first and last frames of interest for the LtR trial as Frame $n_{\text{first}}^{\text{LtR}} = l_{m1}^{\mathrm{S}_1}$ and $n_{\text{last}}^{\text{LtR}} = l_{\text{first}}^{\mathrm{S}_2}$, respectively, and for the RtL trial as $n_{\text{first}}^{\text{RtL}} = l_{\text{last}}^{\mathrm{S}_2}$ and $n_{\text{last}}^{\text{RtL}} = l_{m2}^{\mathrm{S}_1}$, respectively. If $\Delta l_{\max}^{\mathrm{S}_1} < \Delta l_{\max}^{\mathrm{S}_2}$, which indicates the trial direction is RtL $\rightarrow$ LtR, then $n_{\text{first}}^{\text{RtL}} = l_{m1}^{\mathrm{S}_2}$, $n_{\text{last}}^{\text{RtL}} = l_{\text{first}}^{\mathrm{S}_1}$, and $n_{\text{first}}^{\text{LtR}} = l_{\text{last}}^{\mathrm{S}_1}$, $n_{\text{last}}^{\text{LtR}} = l_{m2}^{\mathrm{S}_2}$. The overall frames-of-interest detection scheme is summarized in Algorithm 1.

In all experiments, we set a size of 360×80 pixels for each frame-segment, and a threshold $\tau = 10000$ for peak detection. For evaluation, we manually label the frames where all three markers on the same leg first and last appear for each trial as the ground truth, and compare them with the corresponding detected frames using the following "frame difference rate (FDR)" measure:

Algorithm 1: Frames-of-interest detection for marker tracking.

Input: $\mathbf{D}$, N, S_1, S_2.
Output: $n_{\text{first}}^{\text{LtR}}$, $n_{\text{last}}^{\text{LtR}}$, $n_{\text{first}}^{\text{RtL}}$, $n_{\text{last}}^{\text{RtL}}$.
initialize $n = 1$;
for $n \leq N$ **do**
  $\mathrm{F}_n = \mathbf{D}(n)$;
  if $n = 1$ **then**
    $\mathrm{F}_1, \mathrm{S}_1, \mathrm{S}_2, \Rightarrow \mathrm{H}_1^x$;
  else
    $\mathrm{F}_n, \mathrm{S}_1, \mathrm{S}_2, \Rightarrow \mathrm{H}_n^x \Rightarrow \Delta\mathrm{H}_n^x \Rightarrow \sum \Delta\mathrm{H}_n^x$;
  $n = n + 1$;
$\{n, \sum \Delta\mathrm{H}_n^x\}$, peak detection $\Rightarrow \mathrm{P}^x$;
P^x, difference of the frame numbers $\Rightarrow \Delta l^x \Rightarrow \Delta l_{\max}^x \Rightarrow$ $\mathrm{P}_1^x,\ \mathrm{P}_2^x$;
if $\Delta l_{\max}^{\mathrm{S}_1} > \Delta l_{\max}^{\mathrm{S}_2}$ **then**
  $\mathrm{P}_1^x,\ \mathrm{P}_2^x,\ \text{LtR} \rightarrow \text{RtL} \Rightarrow n_{\text{first}}^{\text{LtR}},\ n_{\text{last}}^{\text{LtR}},\ n_{\text{first}}^{\text{RtL}},\ n_{\text{last}}^{\text{RtL}}$;
else
  $\mathrm{P}_1^x,\ \mathrm{P}_2^x,\ \text{RtL} \rightarrow \text{LtR} \Rightarrow n_{\text{first}}^{\text{RtL}},\ n_{\text{last}}^{\text{RtL}},\ n_{\text{first}}^{\text{LtR}},\ n_{\text{last}}^{\text{LtR}}$;

$$\text{FDR} = \frac{|n_{\text{detected}} - n_{\text{labelled}}|}{n_{\text{labelled}}} \times 100\% \tag{1}$$

where n_{labelled} and n_{detected} denote the frame number of the manually labelled frame and detected frame, respectively.

2.4 Marker tracking and autonomous knee angle calculation

The marker tracking process is initialized by the marker-template selection for hip, knee, and ankle markers via a mouse-click in the detected first frame of interest using a bespoke GUI. Three markers are individually tracked via Structural-Similarity (SSIM) [37] with a template-matching motion search scheme, within a Search Area (SA) whose position and size are determined by a discrete Kalman filter (DKF) [36]. The centre coordinate of each tracked marker in each frame is simultaneously determined for autonomous knee angle calculation.

In particular, for each marker, we first define the SA of size $h \times h$ pixels, and the centre coordinate and velocity given by $\hat{\mathbf{s}}_i = [f_i\ g_i\ v_i f_i\ v_i g_i]^T$, where f_i and g_i denote the column and row centre coordinate of the SA in frame F_i, respectively, and v_i denotes velocity. The column and row centre coordinate of the marker in F_i are denoted as $\hat{\mathbf{u}}_i = [c_i\ r_i]^T$. We adopt a DKF [36], whose dynamic and observation models are constructed by $\hat{\mathbf{s}}_i$ and $\hat{\mathbf{u}}_i$, respectively. This DKF consists of a prediction phase and a correction phase:

$$\begin{aligned}
&\text{prediction phase}: \begin{cases} \hat{\mathbf{s}}_i^- = \mathbf{R}\hat{\mathbf{s}}_{i-1}, \\ \mathbf{X}_i^- = \mathbf{R}\mathbf{X}_{i-1}\mathbf{R}^T + \mathbf{B}. \end{cases} \\
&\text{correction phase}: \begin{cases} \mathbf{K}_i = \mathbf{X}_i^-\mathbf{Q}^T(\mathbf{Q}\mathbf{X}_i^-\mathbf{Q}^T + \mathbf{E})^{-1} \\ \hat{\mathbf{s}}_i = \hat{\mathbf{s}}_i^- + \mathbf{K}_i(\hat{\mathbf{u}}_i - \mathbf{Q}\hat{\mathbf{s}}_i^-) \\ \mathbf{X}_i = (1 - \mathbf{K}_i\mathbf{Q})\mathbf{X}_i^- \end{cases}
\end{aligned} \tag{2}$$

where $\hat{\mathbf{s}}_i^-$ is the a priori estimate of $\hat{\mathbf{s}}_i$ in F_i, $\hat{\mathbf{s}}_{i-1}$ is the a posteriori estimate, $\mathbf{R}$ is the state transition matrix, $\mathbf{X}_i^-$ is the a posteriori covariance matrix, $\mathbf{B}$ is the process noise covariance matrix pre-computed by running the filter off-line based on the assumption that $\mathbf{B}$ is time invariant [36], $\mathbf{K}_i$ is the Kalman gain, $\mathbf{Q}$ is the observation matrix, and $\mathbf{E}$ is the measurement error covariance matrix pre-computed by running the filter off-line based on the assumption that $\mathbf{E}$ is constant across all frames [36]; v_i, which is in $\hat{\mathbf{s}}_i$, is determined by the DKF (K-velocity). See [36] for calculation of $\mathbf{R}$, $\mathbf{X}_i^-$, and $\mathbf{Q}$.

The above filter is initialized by $\hat{\mathbf{s}}_1^- = [f_0\ g_0\ 0\ 0]^T$ and $\hat{\mathbf{u}}_1 = [c_1\ r_1]^T$, where $f_0 = c_1$, $g_0 = r_1$, and (c_1, r_1) denotes the centre coordinate of the marker-template. The four edges of the SA are dynamically updated in each frame based on $v_{i-1}f_{i-1}$ and $v_{i-1}g_{i-1}$. Given the duration of one frame is t seconds, if $v_{i-1}f_{i-1} \geq 0$, the right edge of $\hat{\mathbf{s}}_i$ is shifted to the right by $v_{i-1}f_{i-1}t$ pixels; otherwise, the left edge of the SA is shifted to the left by the same number of pixels. Similarly, if $v_{i-1}g_{i-1} \geq 0$, the bottom edge of the SA is shifted down by $v_{i-1}g_{i-1}t$ pixels; otherwise, the top edge of the SA is shifted up by the same number of pixels.

For template matching, we adopt SSIM with a motion full-search scheme to track the marker within the SA. SSIM is an image quality assessment metric shown in Equation (3) which combines the luminance (L), contrast (C), and structure (J) comparison between a candidate block within the updated SA ($\mathrm{SA}_{\text{updated}}$) in F_n, denoted as $\mathbf{a}_n$, and its corresponding marker-template, denoted as $\mathbf{b}$, where $0 < \text{SSIM}(\mathbf{a}_n, \mathbf{b}) \leq 1$. The candidate block with the largest $\text{SSIM}(\mathbf{a}_n, \mathbf{b})$, over all $\mathbf{a}_n$ in SA denoted as $\mathbf{a}_n^{\text{best}}$, is designated as the tracked marker; we denote its centre coordinate as $\hat{\mathbf{u}}_{n+1}$, which is used to update the observation and dynamic models in the above Kalman filter.

$$\text{SSIM}(\mathbf{a}_n, \mathbf{b}) = [L(\mathbf{a}_n, \mathbf{b})] \cdot [C(\mathbf{a}_n, \mathbf{b})] \cdot [J(\mathbf{a}_n, \mathbf{b})]. \tag{3}$$

There exist several occlusion phases (OP) for the hip marker due to arm swing. We address this occlusion problem by setting a heuristically determined threshold τ_{op}, that is, the frame where $\text{SSIM}(\mathbf{a}_n^{\text{best}}, \mathbf{b}) \leq \tau_{op}$ is the first frame of occlusion, and its frame number is denoted by $n_{\text{start}}^{\text{OP}}$. The SSIM exhaustive search algorithm continues to process the subsequent frames until $\text{SSIM}(\mathbf{a}_n^{\text{best}}, \mathbf{b}) > \tau_{op}$, which indicates that the hip marker has appeared again after occlusion, and its frame number is denoted by $n_{\text{end}}^{\text{OP}}$. Next, non-linear interpolation, based on the centre coordinates of the hip marker and the distances between the hip and knee markers in $\mathrm{F}_{n_{\text{start}}^{\text{OP}}}$ and $\mathrm{F}_{n_{\text{end}}^{\text{OP}}}$, is performed to estimate the centre coordinates of the hip marker, denoted as $\{\hat{\mathbf{u}}\}_{\text{OP}}$, within the occluded frames $\{\mathrm{F}_{n_{\text{start}}^{\text{OP}}}, \ldots, \mathrm{F}_{n_{\text{end}}^{\text{OP}}}\}$. The overall marker tracking procedure for each marker is summarized in Algorithm 2.

We note that, the SSIM threshold τ_{op} was heuristically chosen by testing the marker tracking component in one training sample video so that our marker tracking component is able to track the markers well with robust occlusion handling. The knee angle data can only be acquired when the marker tracking component is robust enough, *i.e.*, tracking the markers without mis-tracking.

Since each video frame contains three channels, (R)ed, (G)reen, and (B)lue, we perform marker tracking in the three channels independently, and then calculate the mean values of the centre coordinates of the tracked marker, that is, $\{\hat{\mathbf{u}}\}^{\text{RGB}} = \frac{1}{3}(\{\hat{\mathbf{u}}\}^{\text{R}} + \{\hat{\mathbf{u}}\}^{\text{G}} + \{\hat{\mathbf{u}}\}^{\text{B}})$. In another approach, the grayscale scheme, we convert each frame into a single grayscale channel before marker tracking, and perform the marker tracking once, getting the centre coordinates of the tracked marker $\{\hat{\mathbf{u}}\}^{\text{grayscale}}$.

The knee angle is automatically calculated during RGB and grayscale marker tracking. Figures 3(a) and (b) show the sample knee angle plots of a stroke patient and a healthy volunteer, respectively, during a walking trial using grayscale marker tracking.

As shown in Figure 4, the overall proposed single-camera gait analysis system process includes video acquisition with camera calibration, autonomous detection of the first and last tracking-frames, marker tracking, and autonomous knee angle calculation.

2.5 Validation against VICON

The system is validated on 15 participants, including 10 stroke patients (7 males and 3 females, age range 48-100 with mean value 69.03 and standard deviation 14.04) recruited between June 2011 and July 2012 from 4 UK hospitals [51], and 5 healthy volunteers (3 males and 2 females, age range 26-35 with mean value 29.20 and standard deviation 4.44) recruited during May 2014 from the University of Strathclyde staff. All recruited stroke patients had hemiplegia. The level of impairment was assessed using the Functional Ambulatory Category (FAC). The patient selection criteria included patients

Algorithm 2: Marker tracking for a sample LtR walking trial.

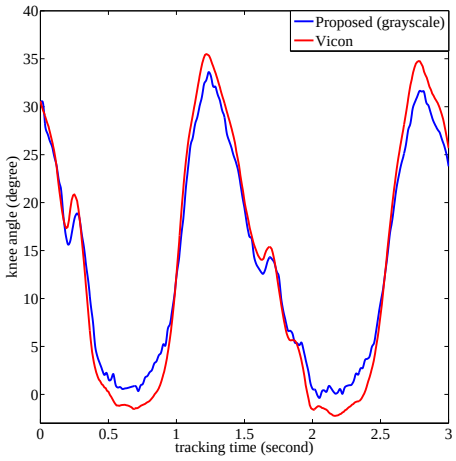
Input: $\mathbf{D}$, $n_{\text{first}}^{\text{LtR}}$, $n_{\text{last}}^{\text{LtR}}$, $\mathbf{b}$, $\mathbf{R}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{E}$, τ_{op}, t.
Output: $\{\hat{\mathbf{u}}\}$.
initialize $\hat{\mathbf{s}}_{n-1}$, $\hat{\mathbf{u}}_{n-1}$, $n = n_{\text{first}}^{\text{LtR}} + 1$,
occlusion state (OS) $= 0$, occlusion phase (OP) $= 0$;
for $n \leq n_{\text{last}}^{\text{LtR}}$ **do**
  $\mathrm{F}_n = \mathbf{D}(n)$;
  if OS $= 0$ **then**
    DKF: $\mathbf{R}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{E}$, $\hat{\mathbf{s}}_{n-1}$, $\hat{\mathbf{u}}_{n-1}$, Equation (2) $\Rightarrow$ $\hat{\mathbf{s}}_{n-1}^{-}$, $\mathbf{S}_n$;
    SSIM: $\mathbf{b}$, F_n, $\mathbf{S}_n$, Equation (3) $\Rightarrow$ $\mathbf{a}_n^{\text{best}}$;
    if SSIM$(\mathbf{a}_n^{\text{best}}, \mathbf{b}) < \tau_{op}$ **then**
      OS $= 1$, OP $=$ OP $+ 1$, $n_{\text{start}}^{\text{OP}} = n$;
      $\overline{v_c} = (\hat{\mathbf{u}}_{n-1}(1) - \hat{\mathbf{u}}_{n_{\text{first}}^{\text{LtR}}}(1))/(n - 1 - n_{\text{first}}^{\text{LtR}})$,
      $\overline{v_r} = (\hat{\mathbf{u}}_{n-1}(2) - \hat{\mathbf{u}}_{n_{\text{first}}^{\text{LtR}}}(2))/(n - 1 - n_{\text{first}}^{\text{LtR}})$;
      $\mathbf{S}_n^{'\text{CC}} = \hat{\mathbf{u}}_{n-1}(1)$, $\mathbf{S}_n^{'\text{RC}} = \hat{\mathbf{u}}_{n-1}(2)$;
    else
      $\mathbf{a}_n^{\text{best}} \Rightarrow \hat{\mathbf{u}}_n$;
  if OS $= 1$ **then**
    SSIM: $\mathbf{b}$, F_n, $\mathbf{S}_n^{'}$, Equation (3) $\Rightarrow$ $\mathbf{a}_n^{\text{best}}$;
    if SSIM$(\mathbf{a}_n^{\text{best}}, \mathbf{b}) \geq \tau_{op}$ **then**
      $\mathbf{a}_n^{\text{best}} \Rightarrow \hat{\mathbf{u}}_n$;
      $n_{\text{end}}^{\text{OP}} = n - 1$, OS $= 0$;
      $\{\mathrm{F}_{n_{\text{start}}^{\text{OP}}}, ..., \mathrm{F}_{n_{\text{end}}^{\text{OP}}}\}, \Rightarrow \{\hat{\mathbf{u}}\}_{\text{OP}}$;
      DKF: $\mathbf{R}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{E}$, $\hat{\mathbf{s}}_{n_{\text{start}}^{\text{OP}}-1}$, $\{\hat{\mathbf{u}}\}_{\text{OP}}$, Equation (2) $\Rightarrow$ $\{\hat{\mathbf{s}}^{-}\}_{\text{OP}}$, $\{\mathbf{S}\}_{\text{OP}}$;
    else
      $\mathbf{S}_n^{'\text{CC}} = \mathbf{S}_n^{'\text{CC}} + \overline{v_c}t$;
      $\mathbf{S}_n^{'\text{RC}} = \mathbf{S}_n^{'\text{RC}} + \overline{v_r}t$;

with a walking ability from FAC score 1 to FAC score 5 but with a) abnormal initial floor contact and/or b) impaired ability to take full body weight through the paretic lower limb in stance. Each participant performs two pairs of LtR and RtL walking trials, each trial includes at least 2 consecutive gait cycles. Thus, the test dataset includes 40 trials for stroke patients and 20 trials for healthy volunteers. The knee angle data is down-sampled from 210fps to 100fps, for a fair comparison against VICON (100fps). The data processing module is implemented in Matlab R2014b on a laptop running Windows 8.1, with Core i7 2820QM 2.3GHz processor and 16GB RAM. In this section, we show the result of the evaluation of frames-of-interest detection, knee angle validation against VICON, and accuracy investigation of RGB and grayscale marker tracking. In all our experiments we set SSIM threshold $\tau_{op} = 0.4$ (see Section 2.4), which gives the best result.

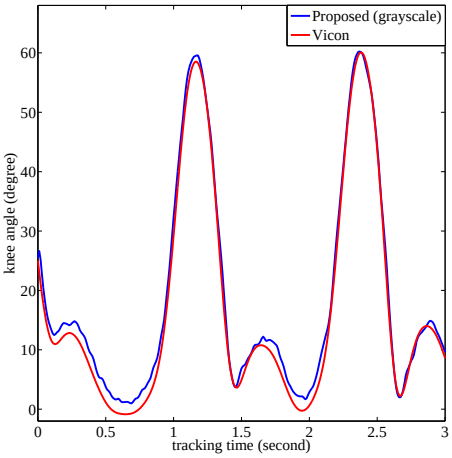
We first group the knee angle dataset for all stroke patients (healthy volunteers) together as a vector W_p, where $p \in \{\text{RGB}, \text{(G)rayscale}, \text{(V)icon}\}$. We then calculate the R-squared value (RSV), max error (ME), and root mean square error (RMSE) between W_{RGB} and W_{V}, and between W_{G} and W_{V}; we use the Bland-Altman plot [52] between W_{RGB} and W_{V}, and between W_{G} and W_{V}, and calculate the mean error (MEE), 95% confidence interval (CI), and linear fit with slope (S) and intercept (I), based on the constructed Bland-Altman plot. Tables 1, 2 and 3 show the knee angle validation result based on both RGB and grayscale marker tracking schemes on paretic legs and non-paretic legs of stroke patients and healthy volunteers, respectively.

2.6 *Performance comparison of RGB and grayscale marker tracking*

We choose the best marker tracking scheme by comparing the performance of the RGB and grayscale marker tracking using the



(a)



(b)

Fig. 3: (a) Sample knee angle of a stroke patient using grayscale marker tracking; (b) Sample knee angle of a healthy volunteer using grayscale marker tracking.

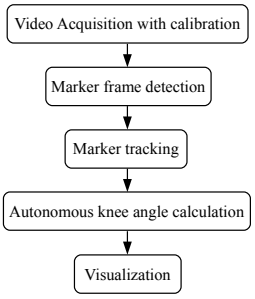


Fig. 4: System workflow.

following comparison metrics:

$$\phi_{\text{RSV}} = \frac{\text{RSV}_{\text{G}} - \text{RSV}_{\text{RGB}}}{\text{RSV}_{\text{G}}}$$

$$\phi_{\text{ME}} = \frac{|\text{ME}_{\text{G}}| - |\text{ME}_{\text{RGB}}|}{|\text{ME}_{\text{G}}|}$$

$$\phi_{\text{RMSE}} = \frac{|\text{RMSE}_{\text{G}}| - |\text{RMSE}_{\text{RGB}}|}{|\text{RMSE}_{\text{G}}|}$$

$$\phi_{\text{MEE}} = \frac{|\text{MEE}_{\text{G}}| - |\text{MEE}_{\text{RGB}}|}{|\text{MEE}_{\text{G}}|}$$

Table 1 Knee angle on patients' paretic legs. BA=Bland-Altman plot.

metric	RGB	grayscale
R-squared value	0.9819	0.9818
max error (deg)	-7.4064	-7.3051
root mean square error (deg)	3.6567	3.6657
BA mean error (deg)	-1.1181	-1.1159
BA 95% confidence interval	(-5.9958, 3.7595)	(-6.0028, 3.7710)
BA linear fit slope	-0.0988	-0.0990
BA linear fit intercept	0.5953	0.6013
average execution time (s)	1318.2641	703.1875

Table 2 Knee angle on patients' non-paretic legs. BA=Bland-Altman plot.

metric	RGB	grayscale
R-squared value	0.9829	0.9827
max error (deg)	-9.1448	-9.3376
root mean square error (deg)	3.3296	3.3591
BA mean error (deg)	-0.7616	-0.7598
BA 95% confidence interval	(-5.5495, 4.0263)	(-5.5720, 4.0524)
BA linear fit slope	-0.0527	-0.0528
BA linear fit intercept	0.3537	0.3565
average execution time (s)	1294.5768	741.2971

Table 3 Knee angle on healthy volunteers. BA=Bland-Altman plot.

metric	RGB	grayscale
R-squared value	0.9709	0.9711
max error (deg)	12.2424	12.2055
root mean square error (deg)	3.9931	4.0129
BA mean error (deg)	2.9804	3.0127
BA 95% confidence interval	(-2.2286, 8.1894)	(-2.1834, 8.2088)
BA linear fit slope	0.0149	0.0162
BA linear fit intercept	2.7778	2.7918
average execution time (s)	1282.6315	697.4723

$$\phi_{\mathrm{CI}} = \frac{(\mathrm{UCI_G} - \mathrm{LCI_G}) - (\mathrm{UCI_{RGB}} - \mathrm{LCI_{RGB}})}{\mathrm{UCI_G} - \mathrm{LCI_G}}$$

$$\phi_{\mathrm{S}} = \frac{|\mathrm{S_G}| - |\mathrm{S_{RGB}}|}{|\mathrm{S_G}|}$$

$$\phi_{\mathrm{I}} = \frac{|\mathrm{I_G}| - |\mathrm{I_{RGB}}|}{|\mathrm{I_G}|}$$

$$\phi_{\mathrm{ET}} = \frac{|\mathrm{ET_G}| - |\mathrm{ET_{RGB}}|}{|\mathrm{ET_G}|}$$

In particular, we find the best performing method (grayscale or RGB) by comparing the errors ϕ obtained by validating these two methods with VICON normalized by the error obtained by validating the grayscale method with VICON. If $\phi > 0$, we conclude that the RGB marker tracking performs better, except for the R-squared value where $\phi > 0$ means that the grayscale method is better.

3 Results

The average time to generate knee-angle results is 30 minutes for each participant, which includes 5 minutes for camera and tripod assembly, 2 minutes for adjustment of the camera height and distance to the participant, 2 minutes for marker attachment, 5 minutes for video recording, and the rest for data processing. The data processing module starts with the autonomous frames-of-interest detection, followed by marker-template selection via a mouse-click, autonomous grayscale marker tracking and knee angle calculation.

3.1 Evaluation of autonomous frames-of-interest detection

The proposed autonomous frames-of-interest detection scheme accurately detects the first and last tracking frames when all markers are present for the marker tracking process in all videos with a mean FDR (see Equation (1)) value of 0.2%, that is, given a 1000-frame video, the difference between n_{detected} and n_{labelled} is only 2 on average. Moreover, the average execution time for each detection process is only 6.2351s. Our autonomous frame-of-interest detection scheme shows negligible FDR, *i.e.*, the proposed frame-of-interest detection method successfully recognizes both the first and last frames when all three markers are present for the marker tracking process – thus it can be used to replace manual labelling.

3.2 Validation against VICON

In general, RGB and grayscale schemes achieve similar accuracy, where the grayscale scheme performs almost 50% faster than the RGB scheme. In particular, as shown in Tables 1, 2 and 3, with the grayscale scheme, the max error of the knee angle validation is -7.3051, -9.3376, and 12.2055 degrees, on patients' paretic leg, non-paretic leg, and for healthy volunteers, respectively. For all validation data, the R-squared value is at least 0.9711, root mean square error is at most 4.0129 degrees, mean error is at most 3.0127 degrees, the 95% confidence range is approximately 10 degrees, the slope and intercept of the linear fit is at most -0.0990 and 2.7918, respectively.

3.3 Performance comparison of RGB and grayscale marker tracking

The comparison results shown in Table 4 indicate that the RGB marker tracking scheme shows performance that is closer to VICON result in terms of 6 out of 8 comparison metrics for stroke patients group and 4 out of 8 for healthy volunteers compared to that of grayscale scheme, where the RGB marker tracking scheme is 8% better than grayscale scheme. However, grayscale scheme shows over 80% faster performance in processing speed (see AET metric in Table 4) than the RGB scheme with negligible performance loss in terms of RSV, ME, RMSE, MEE, 95% CI, and the intercept of the linear fit.

4 Discussion

In this section, we discuss the system operation speed, performance, potential applications and improvements.

First, as shown in Section 3.1, the proposed frame-of-interest detection method achieves high detection rate and thus can be used to replace manual labelling. Our histogram-based FOI method is indeed motivated by background subtraction. Background subtraction returns a foreground mask given (1) a subtraction image of the current frame and the reference frame and (2) a heuristically set pixel intensity threshold. That is, background subtraction relies on pixel intensity subtraction and a single pixel intensity threshold. However, our histogram-based FOI detection method returns the total number of the difference of the number of pixels in each histogram quantization bin given (1) histograms of the current frame and the reference frame and (2) a heuristically set number of pixels threshold. That is, our method relies on pixel number difference and a single pixel number threshold. Unlike background subtraction that requires additional steps (*e.g.*, count the number of non-zero pixels in the foreground mask) for FOI detection, the output of our method, the total number of the difference of the number of pixels in each histogram quantization bin, can be directly used to determine FOI. On the other hand, the recent neural-network-based methods, *e.g.*, [53] and [54], could provide more reliable results, however, both [53] and [54] require large amount of data collection (*i.e.*, photos or video sequences that is similar to the laboratory environment in our experimentation) for model training. Since data collection and data labelling require significant amount of time and human effort, it is impractical to apply such neural-network-based methods to our system.

Table 4 Result of performance comparison between RGB and grayscale marker tracking. BA=Bland-Altman plot.

	stroke patients		healthy volunteers	
	ϕ	winner	ϕ	winner
R-squared value	ϕ_{RSV}=-0.0204%	RGB	ϕ_{RSV}=0.0206%	grayscale
max error (deg)	ϕ_{ME}=2.0648%	RGB	ϕ_{ME}=-0.3023%	G
root mean square error (deg)	ϕ_{RMSE}=0.2717%	RGB	ϕ_{RMSE}=0.4934%	RGB
BA mean error (deg)	ϕ_{ME}=-0.2139%	grayscale	ϕ_{ME}=1.0721%	RGB
BA 95% confidence interval	ϕ_{CI}=0.3435%	RGB	ϕ_{CI}=-0.2483%	grayscale
BA linear fit slope	ϕ_{S}=0.1433%	RGB	ϕ_{S}=8.0247%	RGB
BA linear fit intercept	ϕ_{I}=1.0212%	RGB	ϕ_{I}=0.5015%	RGB
average execution time (s)	ϕ_{ET}=-80.1331%	grayscale	ϕ_{ET}=-83.8971%	grayscale

Next, the proposed marker tracking scheme is more reliable than the state-of-the-art object tracking methods of [55] and [56], as shown in our conference paper [34], where our tracking system significantly outperforms the state-of-the-art tracking methods with at least 33.2% improvement in marker detection rate. As consistent in validation studies such as [57], [58], and [32], a sample size of 15 is used for knee angle validation; the knee angle validation against VICON shows good agreement for all stroke patients and healthy volunteers. Thus the proposed system is robust to stochastic and sudden movements of stroke patients. The efficient performance of the grayscale marker tracking scheme indicates that it is sufficient to convert all RGB frames into grayscale, and then perform tracking and processing on grayscale frames, with the advantage of faster processing time. Again, the neural-network-based methods, *e.g.*, [59] and [54], require large amount of data collection for model training. Since data collection (walking trials of multiple healthy subjects and patients) and data labelling (human anatomical landmarks that are required for human motion tracking) require significant amount of time and human effort, it is impractical to apply such neural-network-based methods to our system. Unlike [59] and [54], our system does not require any training data, while achieving comparable performance compared to the gold standard VICON optical motion analysis system, as shown in Section 3.

There are two types of errors. The first type of error is caused by the deviation between the knee angle plane and the camera scene plane. The second type of error originates from the fundamental difference in defining the joint centres: (1) As explained in Section 2.2, both bulls-eye and retroreflective markers are attached onto the knee and ankle joints. Although such marker attachment does not affect our single-camera and VICON gait analysis systems according to the experimental results, the joint centres of our system are determined solely on the 2D spatial location of the bulls-eye markers in the video sequences, whereas the joint centres of VICON are determined by the Plug-in-Gait protocol [50]. (2) Hip joint centre (HJC) between the gold standard VICON 3D system and our proposed 2D system. That is, the former uses the regression equation decribed by Davis et al. [50] to calculate the HJC location for 3D kinematics, whereas the latter places the marker on the head of the greater trochanter for HJC calculation. Note that the knee angle validation against VICON on study patients' paretic leg shows worse result compared to that on study patients' non-paretic leg for all comparison criteria. This may be due to the irregular gait patterns performed by the study patient's paretic leg where the deviation between the knee angle plane and the camera scene plane is larger than that performed by the patient's non-paretic leg. Note also that the Bland-Altman plot mean error of stroke patients (Tables 1 and 2) is approximately 2 degrees smaller in amplitude compared to healthy volunteers (Table 3), with a 95% confidence interval approximately 0.7 degree smaller. This is due to: (1) the knee range of motion in healthy volunteers being greater than stroke patients given the fact that stroke patients generally perform synergistic gait pattern during walking [60, 61] while healthy individuals perform selective joint movements [62]. This difference in knee range of motion occurs as a result of the deviation between the knee angle plane and the camera scene plane for healthy volunteers being larger than that for stroke patients; (2) both groups have small sample size, and the healthy group was half the size of the patient group. Lastly our 2D portable single-camera gait system achieved a 95% confidence interval of about 10 degrees for both groups compared to the clinically acceptable level of error for 3D kinematics, which is 5 degree [63]. We stress that our proposed 2D system is a much cheaper option and less time consuming, albeit at the small cost of sacrificing a modest amount of accuracy compared to traditional optical motion analysis systems.

5 Conclusion

Emerging 2D single-camera systems are cost effective and highly portable with adequate fidelity of gait parameters compared to laboratory-based motion analysis systems. We proposed a portable single-camera gait analysis system with autonomous frames-of-interest detection and grayscale marker tracking functionality. The proposed system is robust to the background in the video capture room and study participant's clothing colours, autonomously tracks markers and calculates the knee angle in contrast to current video analysis software such as Pro-Trainer and Siliconcoach. Experimental results show that the proposed system can automatically detect the frames-of-interest and measure the knee angle. Furthermore, we show that, for all measures of assessment, use of grayscale video incurs a very small performance loss compared to when RGB video is used, but considerably reduces processing time. Future work consists of extending the system to a stereo 2D-camera system or a single depth sensing device, and investigation of the error with a focus on marker trajectories.

Our future work will be focused on further improvement of performance and potential measurement capability of more gait parameters using a stereo 2D-camera system or a single depth sensing device [2, 13–19] with a high frame rate, without sacrificing portability, to remove the parallax error, and leverage the 3D information for quantifying a larger number of gait parameters such as hip, knee, and ankle angles in both the sagittal and frontal planes, and pelvis tilt, calculating temporal-spatial parameters (step length and width, stride length, step time, cadence and step length symmetry), gait speed and measuring sagittal / frontal plane knee motion, but at an increased processing complexity. Furthermore, neural-network-based methods are what we could consider in future research that involve evaluations with public availability of large labelled datasets of walking trials.

6 Acknowledgements

We confirm that any aspect of the study covered in this manuscript that has involved stroke patients has been conducted with ethical approval of both National Health Service (NHS) and University of Strathclyde, and healthy volunteers with the ethical approval of University of Strathclyde. All 15 participants, including 10 stroke patients and 5 healthy volunteers, read the participant information sheets and completed the consent forms before data collection. The authors would like to greatly appreciate the effort of all participants who were recruited to validate the system.

Competing interests: None declared.

Funding: None.

Ethical approval was given on 17-Apr-2014. The relevant Judgement's reference number is DEC.BioMed.2014.33.

7 References

1 Langhorne, P., Coupar, F., Pollock, A.: 'Motor recovery after stroke: a systematic review', The Lancet Neurology, 2009, **8**, (8), pp. 741–754
2 Muro-de-la Herran, A., Garcia.Zapirain, B., Mendez.Zorrilla, A.: 'Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications', Sensors, 2014, **14**, (2), pp. 3362–3394
3 Toro, B., Nester, C., Farren, P.: 'A review of observational gait assessment in clinical practice', Physiotherapy Theory and Practice, 2003, **19**, pp. 137–149
4 Ferrarello, F., Bianchi, V.A.M., Baccini, M., Rubbieri, G., Mossello, E., Cavallini, M.C., et al.: 'Tools for observational gait analysis in patients with stroke: a systematic review', Physical Therapy, 2013, **93**, (12), pp. 1673–1685
5 Li, G., Liu, T., Yi, J., Wang, H., Li, J., Inoue, Y.: 'The lower limbs kinematics analysis by wearable sensor shoes', IEEE Sensors Journal, 2016, **16**, (8), pp. 2627–2638
6 Altaf, M.U.B., Butko, T., Juang, B.H.: 'Acoustic gaits: Gait analysis with footstep sounds', IEEE Transactions on Biomedical Engineering, 2015, **62**, (8), pp. 2001–2011
7 Richards, J.G.: 'The measurement of human motion: A comparison of commercially available systems', Human Movement Science, 1999, **18**, (5), pp. 589–602
8 Hagler, S., Austin, D., Hayes, T.L., Kaye, J., Pavel, M.: 'Unobtrusive and ubiquitous in-home monitoring: a methodology for continuous assessment of gait velocity in elders', IEEE Transactions on Biomedical Engineering, 2010, **57**, (4), pp. 813–820
9 Soda, P., Carta, A., Formica, D., Guglielmelli, E. 'A low-cost video-based tool for clinical gait analysis'. In: Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE. (Minneapolis, MN, 2009. pp. 3979–3982
10 Yang, M., Zheng, H., Wang, H., McClean, S., Newell, D.: 'igait: An interactive accelerometer based gait analysis system', Computer Methods and Programs in Biomedicine, 2012, **108**, (2), pp. 715–723
11 Ugbolue, U.C., Papi, E., Kaliarntas, K.T., Kerr, A., Earl, L., Pomeroy, V.M., et al.: 'The evaluation of an inexpensive, 2D, video based gait assessment system for clinical use', Gait & Posture, 2013, **38**, (3), pp. 483–489
12 Wall, J.C., Devlin, J., Khirchof, R., Lackey, B.: 'Measurement of step widths and step lengths: a comparison of measurements made directly from a grid with those made from a video recording', Journal of Orthopaedic & Sports Physical Therapy, 2000, **30**, (7), pp. 410–417
13 Hondori, H.M., Khademi, M.: 'A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation', Journal of Medical Engineering, 2014, **2014**, pp. 1–16
14 Ye, M., Yang, C., Stankovic, V., Stankovic, L., Kerr, A.: 'Kinematics analysis multimedia system for rehabilitation', International Conference on Image Analysis and Processing, 2015, pp. 571–579
15 Ye, M., Yang, C., Stankovic, V., Stankovic, L., Kerr, A.: 'Gait analysis using a single depth camera', IEEE Global Conference on Signal and Information Processing, 2015, pp. 285–289
16 Ye, M., Yang, C., Stankovic, V., Stankovic, L., Kerr, A.: 'A depth camera motion analysis framework for tele-rehabilitation: Motion capture and person-centric kinematics analysis', IEEE Journal of Selected Topics in Signal Processing, 2016, **10**, (5), pp. 877–887
17 Ye, M., Yang, C., Stankovic, V., Stankovic, L., Cheng, S.: 'Gait phase classification for in-home gait assessment', IEEE International Conference on Multimedia and Expo, 2017, pp. 1524–1529
18 Gholami, F., Trojan, D.A., KÃűvecses, J., Haddad, W.M., Gholami, B.: 'A microsoft kinect-based point-of-care gait assessment framework for multiple sclerosis patients', IEEE Journal of Biomedical and Health Informatics, 2017, **21**, (5), pp. 1376–1385
19 Lim, C.D., Wang, C.M., Cheng, C.Y., Chao, Y., Tseng, S.H., Fu, L.C.: 'Sensory cues guided rehabilitation robotic walker realized by depth image-based gait analysis', IEEE Transactions on Automation Science and Engineering, 2016, **13**, (1), pp. 171–180
20 Farah, J.D., Baddour, N., Lemaire, E.D. 'Gait phase detection from thigh kinematics using machine learning techniques'. In: Medical Measurements and Applications (MeMeA), 2017 IEEE International Symposium on. (, 2017. pp. 263–268
21 Gianaria, E., Grangetto, M., Roppolo, M., Mulasso, A., Rabaglietti, E. 'Kinect-based gait analysis for automatic frailty syndrome assessment'. In: IEEE International Conference on Image Processing. (, 2016. pp. 1314–1318
22 Caldas, R., Mundt, M., Potthast, W., de Lima.Neto, F.B., Markert, B.: 'A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms', Gait & Posture, 2017, **57**, pp. 204 – 210
23 Ameli, S., Naghdy, F., Stirling, D., Naghdy, G., Aghmesheh, M.: 'Objective clinical gait analysis using inertial sensors and six minute walking test', Pattern Recognition, 2017, **63**, pp. 246 – 257
24 Tunca, C., Pehlivan, N., Ak, N., Arnrich, B., Salur, G., Ersoy, C.: 'Inertial sensor-based robust gait analysis in non-hospital settings for neurological disorders', Sensors, 2017, **17**, pp. 1 – 29
25 Song, M., Kim, J.: 'An ambulatory gait monitoring system with activity classification and gait parameter calculation based on a single foot inertial sensor', IEEE Transactions on Biomedical Engineering, 2018, **65**, (4), pp. 885–893
26 Bartlett, H.L., Goldfarb, M.: 'A phase variable approach for imu-based locomotion activity recognition', IEEE Transactions on Biomedical Engineering, 2018, pp. 1–1. early access
27 Johansson, T., Wild, C.: 'Telerehabilitation in stroke care — a systematic review', Journal of Telemedicine and Telecare, 2011, **17**, (1), pp. 1–6
28 Laver, K.E., Schoene, D., Crotty, M., George, S., Lannin, N.A., Sherrington, C.: 'Telerehabilitation services for stroke', Cochrane Database of Systematic Reviews, 2013, **12**, pp. CD010255
29 Theodoros, D., Russell, T.: 'Telerehabilitation: Current perspectives', Studies in Health Technology and Informatics, 2008, **131**, pp. 191–209
30 Brennan, D.M., Mawson, S., Brownsell, S.: 'Telerehabilitation: Enabling the remote delivery of healthcare, rehabilitation, and self management', Studies in Health Technology and Informatics, 2009, **145**, pp. 231–248
31 Gregory, P., Alexander, J., Satinsky, J.: 'Clinical telerehabilitation: Applications for physiatrists', PM&R: The journal of injury, function, and rehabilitation, 2011, **3**, (7), pp. 647–656
32 Deltombe, T., Detrembleur, C., Gruwez, G.: 'Comparison of tracker 2-d video software and vicon 3-d system in knee and ankle gait kinematic analysis of spastic patients', Annals of Physical and Rehabilitation Medicine, 2017, **60**, pp. e51
33 Yang, C., Ugbolue, U.C., Carse, B., Stankovic, V., Stankovic, L., Rowe, P.J.: 'Multiple marker tracking in a single-camera system for gait analysis', Congress of the International Society of Biomechanics, 2013, pp. 1–2
34 Yang, C., Ugbolue, U.C., Carse, B., Stankovic, V., Stankovic, L., Rowe, P.J.: 'Multiple marker tracking in a single-camera system for gait analysis', IEEE International Conference on Image Processing, 2013, pp. 3128–3131
35 Kalman, R.E.: 'A new approach to linear filtering and prediction problems', Journal of Fluids Engineering-Transactions of the ASME, 1960, **82**, (1), pp. 35–45
36 Welch, G., Bishop, G. 'An introduction to the kalman filter'. (Chapel Hill, NC, USA, 1995.
37 Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: 'Image quality assessment: from error visibility to structural similarity', IEEE Transactions on Image Processing, 2004, **13**, (4), pp. 600–612
38 Yang, C., Ugbolue, U.C., Kerr, A., Stankovic, V., Stankovic, L., Carse, B., et al.: 'Autonomous gait event detection with portable single-camera gait analysis system', Journal of Sensors, 2016, **2016**, pp. 1–8
39 Kerrigan, D.C., Schaufele, M., Wen, M.N. 'Gait analysis'. In: Delisa, J.A., Gans, B.M., editors. Rehabilitation Medicine: Principles and Practice. 3rd ed. (Philadelphia: Lippincott-Raven Publishers, 1998. pp. 167–187
40 Yang, C., Kerr, A., Stankovic, V., Stankovic, L., Rowe, P., Cheng, S.: 'Human upper limb motion analysis for post-stroke impairment assessment using video analytics', IEEE Access, 2016, **4**, pp. 650–659
41 Chandra Prakash and Kanika Gupta and Anshul Mittal and Rajesh Kumar and Vijay Laxmi: 'Passive Marker Based Optical System for Gait Kinematics for Lower Extremity', Procedia Computer Science, 2015, **45**, pp. 176 – 185
42 McDonald, D.A., Delgadillo, J.Q., Fredericson, M., McConnell, J., Hodgins, M., Besier, T.F.: 'Reliability and accuracy of a video analysis protocol to assess core ability', PM&R: The journal of injury, function, and rehabilitation, 2011, **3**, (3), pp. 204–211
43 Richardson, S., Cooper, A., Alghamdi, G., Alghamdi, M., Altowaijri, A.: 'Assessing knee hyperextension in patients after stroke: comparing clinical observation and Siliconcoach software', International Journal of Therapy and Rehabilitation, 2012, **19**, (3), pp. 163–168
44 Meadows, C.B. 'The influence of polypropylene ankle-foot orthoses on the gait of cerebral palsied children'. University of Strathclyde, 1984
45 Carse, B., Bowers, R., Meadows, B.C., Rowe, P.: 'The immediate effects of fitting and tuning solid ankle–foot orthoses in early stroke rehabilitation', Prosthetics and Orthotics International, 2015, **39**, (6), pp. 454–462
46 Owen, E.: 'The importance of being earnest about shank and thigh kinematics especially when using ankle-foot orthoses', Prosthetics and Orthotics International, 2010, **34**, (3), pp. 254–269
47 Zhang, Z.: 'A flexible new technique for camera calibration', IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, **22**, (11), pp. 1330–1334
48 Pomeroy, V.M., Rowe, P.J., Baron, J.C., Clark, A., Sealy, R., Ugbolue, U.C., et al.: 'The SWIFT Cast trial protocol: A randomized controlled evaluation of the efficacy of an ankle-foot cast on walking recovery early after stroke and the neural-biomechanical correlates of response', International Journal of Stroke, 2012, **7**, (1), pp. 86–93
49 Papi, E., Ugbolue, U.C., Solomonidis, S., Rowe, P.J.: 'Comparative study of a newly cluster based method for gait analysis and plug-in gait protocol', Gait & Posture, 2014, **39**, (Supplement 1), pp. S9–S10
50 Davis.III, R.B., Õunpuu, S., Tyburski, D., Gage, J.R.: 'A gait analysis data collection and reduction technique', Human Movement Science, 1991, **10**, (5), pp. 575–587
51 Pomeroy, V.M., Rowe, P., Clark, A., Walker, A., Kerr, A., Chandler, E., et al.: 'A randomized controlled evaluation of the efficacy of an ankle-foot cast on walking recovery early after stroke swift cast trial', Neurorehabilitation and Neural Repair, 2016, **30**, (1), pp. 40–48
52 Bland, J.M., Altman, D.G.: 'Statistical methods for assessing agreement between two methods of clinical measurement', The Lancet, 1986, **327**, (8476), pp. 307–310
53 G. Lin and A. Milan and C. Shen and I. Reid. 'RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation'. In: IEEE Conference on Computer Vision and Pattern Recognition. (, 2017. pp. 5168–5177
54 K. He and G. Gkioxari and P. DollÃąr and R. Girshick. 'Mask R-CNN'. In: IEEE International Conference on Computer Vision. (, 2017. pp. 2980–2988
55 Ning, J., Zhang, L., Zhang, D., Wu, C.: 'Robust object tracking using joint color-texture histogram', International Journal of Pattern Recognition and Artificial Intelligence, 2009, **23**, (7), pp. 1245–1263
56 Kalal, Z., Mikolajczyk, K., Matas, J.: 'Tracking-learning-detection', IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, **34**, (7), pp. 1409–1422

57 Pfister, A., West, A.M., Bronner, S., Noah, J.A.: 'Comparative abilities of microsoft kinect and vicon 3d motion capture for gait analysis', Journal of Medical Engineering & Technology, 2014, **38**, (5), pp. 274–280
58 Leardini, Alberto and Lullini, Giada and Giannini, Sandro and Berti, Lisa and Ortolani, Maurizio and Caravaggi, Paolo: 'Validation of the angular measurements of a new inertial-measurement-unit based rehabilitation system: comparison with state-of-the-art gait analysis', Journal of NeuroEngineering and Rehabilitation, 2014, **11**, (1), pp. 136
59 Z. Cao and T. Simon and S. Wei and Y. Sheikh. 'Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields'. In: IEEE Conference on Computer Vision and Pattern Recognition. (, 2017. pp. 1302–1310
60 Mazuquin, B.F., Batista JP Jr. , Pereira, L.M., Dias, J.M., Silva, M.F., Carregaro, R.L., et al.: 'Kinematic gait analysis using inertial sensors with subjects after stroke in two different arteries', Journal of Physical Therapy Science, 2014, **26**, (8), pp. 1307–1311
61 Chen, C.L., Chen, H.C., Tang, S.F., Wu, C.Y., Cheng, P.T., Hong, W.H.: 'Gait performance with compensatory adaptations in stroke patients with different degrees of motor recovery', American Journal of Physical Medicine & Rehabilitation, 2003, **82**, (12), pp. 925–35
62 Chaler, J., MÃijller, B., Maiques, A., Pujol, E.: 'Suspected feigned knee extensor weakness: Usefulness of 3d gait analysis. case report', Gait & Posture, 2010, **32**, (3), pp. 354 – 357
63 McGinley, J.L., Baker, R., Wolfe, R., Morris, M.E.: 'The reliability of three-dimensional kinematic gait measurements: a systematic review', Gait & Posture, 2009, **29**, (3), pp. 360–369